Please replace the paragraphs beginning on page 2, line 24, through page 5, line 9, with the following:

Regulation of AAV genes is complex and involves positive and negative regulation of viral transcription. For example, the regulatory proteins Rep 78 and Rep 68 interact with viral promoters to establish a feedback loop (Beaton *et al.* (1989) *J. Virol 63*:4450-4454; Hermonat (1994) *Cancer Lett 81*:129-136). Expression from the p5 and p19 promoters is negatively regulated in *trans* by these proteins. Rep 78 and 68, which are required for this regulation, bind to inverted terminal repeats (ITRs; Ashktorab *et al.* (1989) *J. Virol. 63*:3034-3039) in a site- and strand-specific manner, *in vivo* and *in vitro*. This binding to ITRs induces a cleavage at the TRS and permits the replication of the hairpin structure, thus, illustrating the Rep helicase and endonuclease activities (Im *et al.* (1990) *Cell 61*:447-457; and Walker *et al.* (1997) *J. Virol. 71*:6996-7004), and the role of these non-structural proteins in the initial steps of DNA replication (Hermonat *et al.* (1984) *J. Virol. 52*:329-339). Rep 52 and 40, the two minor forms of the Rep proteins, do not bind to ITRs and are dispensable for viral DNA replication and site-specific integration (Im *et al.* (1992) *J. Virol. 66*:1119-112834; Ni *et al.* (1994) *J. Virol. 68*:1128-1138.

The genome (see, FIG. 1) is organized into two open reading frames (ORFs, designated left and right) that encode structural capsid proteins (Cap) and non-structural proteins (Rep). There are three promoters: p5 (from nucleotides 255 to 261: TATTTAA), p19 (from nucleotide 843 to 849: TATTTAA) and p40 (from nucleotides 1822 to 1827: ATATAA). The right-side ORF (see FIG. 1) encodes three capsid structural proteins (Vp 1-3). These three proteins, which are encoded by overlapping DNA, result from differential splicing and the use of an unusual initiator codon (Cassinoti *et al.* (1988) *Virology 167*:176-184). Expression of the capsid genes is regulated by the p40 promoter. Capsid proteins VP1, VP2 and VP3 initiate from the p40 promoter. VP1 uses an alternate splice acceptor at nucleotide 2201; whereas VP2 and VP3 are derived from the same transcription unit, but VP2 use an ACG triplet as an initiation codon upstream from the start of VP3. On the left side of the genome, two promoters p5 and p19 direct expression of four regulatory

a2

proteins. The left flanking sequence also uses a differential splicing mechanism (Mendelson *et al.* (1986) *J. Virol 60*:823-832) to encode the Rep proteins, designated Rep 78, 68, 52 and 40 on the basis molecular weight. Rep 78 and 68 are translated from a transcript produced from the p5 promoter and are produced from the unspliced and spliced form, respectively, of the transcript. Rep 52 and 40 are the translation products of unspliced and spliced transcripts from the p19 promoter.

AAV and rAAV have many applications, including use as a gene transfer vector, for introducing heterologous nucleic acid into cells and for genetic therapy. Advances in the production of high-titer rAAV stocks to the transition to human clinical trials have been made, but improvement of rAAV production will be complemented with special attention to clinical applications of rAAV vectors as a successful gene therapy approach. Productivity of rAAV (i.e. the amount of vector particles that can be obtained per unitary manufacturing peration) is one of the rate limiting steps in the further development of rAAV as a gene therapy vector. Methods for high throughput production and screening of rAAV have been developed (see, *e.g.*, Drittanti *et al.* (2000) *Gene Therapy 7*:924-929). Briefly, as with the other steps in methods provided herein, the plasmid preparation, transfection, virus productivity and titer and biological activity assessment are intended to be performed in an automatable high throughput format, such as in a 96 well or loci formats (or other number of wells or multiples of 96, such as 384, 1536 . . . 9600, 9984 . . well or loci formats).

A2

SUMMARY

Mutant AAV Rep proteins, nucleic acid molecules encoding such proteins, and rAAV that encode the proteins are provided. Among the rep proteins are

those that result in increased rAAV production in rAAV that encode such mutants, thereby, among a variety of advantages, offering a solution to the need in the gene therapy industry to increase the production of therapeutic vectors without up-scaling manufacturing. Methods of gene therapy using the rAAV are provided.

Directed evolution methods provided in co-pending U.S. provisional application Serial No. 60/315,382, filed as U.S.S.N. 10/022,249, and described herein have been used to identify amino acid "hit" positions in adeno-associated virus (AAV) rep proteins that are relevant for AAV or rAAV production. Those amino acid positions are selected such that a change in the amino acid leads to a change in protein activity either to lower activity or to higher activity compared to native-sequence Rep proteins. The hit positions were then used to generate further mutants designated "leads." Provided herein are the resulting mutant rep proteins that result in either higher or lower levels of AAV or rAAV virus compared to the wild-type (native) Rep protein(s). Nucleic acid molecules that encode the mutant Rep proteins are also provided.

A2 cont'd

Please replace the paragraph on page 5, lines 20-22, with the following:

A3

FIGURE 1 shows the genetic map of AAV, including the location of promoters, and transcripts; amino acid 1 of the Rep 78 gene is at nucleotide 321 in the AAV-2 genome.

Please replace the paragraphs on page 6, lines 21-32, with the following:

As used herein, directed evolution refers to methods that adapt natural proteins or protein domains to work in new chemical or biological environments and/or to elicit new functions. It is more a more broad-based technology than DNA shuffling.

A4

As used herein, high-throughput screening (HTS) refers to processes that test a large number of samples, such as samples of test proteins or cells containing nucleic acids encoding the proteins of interest to identify structures of interest or to identify test compounds that interact with the variant proteins

Af

or cells containing them. HTS operations are amenable to automation and are typically computerized to handle sample preparation, assay procedures and the subsequent processing of large volumes of data.

Please replace the paragraphs on page 7, lines 4-31, with the following:

As used herein, "hits" are mutant proteins that have an alteration in any attribute, chemical, physical or biological property in which such alteration is sought. In the methods herein, hits are generally generated by systematically replacing each amino acid in a protein or a domain thereof with a selected amino acid, typically Alanine, Glycine, Serine or any amino acid, as long as each residue is replaced with the same residue. Hits may be generated by other methods known to those of skill in the art tested by the high-throughput methods herein. For purposes herein a Hit typically has activity with respect to the function of interest that differs by at least 10%, 20%, 30% or more from the wild type or native protein. The desired alteration, which is generally a reduction in activity, will depend upon the function or property of interest.

As used herein, "leads" are "hits" whose activity has been optimized for the particular attribute, chemical, physical or biological property. In the methods herein, leads are generally produced by systematically replacing the hit loci with all remaining 18 amino acids, and identifying those among the resulting proteins that have a desired activity. The leads may be further optimized by replacement of a plurality of "hit" residues. Leads may be generated by other methods known to those of skill in the art and tested by the highthroughput methods herein. For purposes herein a lead typically has activity with respect to the function of interest that differs from the native activity, by a desired amount and by at least 10%, 20%, 30% or more from the wild type or native protein. Generally a Lead will have an activity that is 2 to 10 or more times the native protein for the activity of interest. As with hits, the change in the activity is dependent upon the activity that is "evolved." The desired alteration will depend upon the function or property of interest.

Please replace the paragraph on page 9, lines 3-19, with the following:

As used herein, adeno-associated virus (AAV) is a defective and non-pathogenic parvovirus that requires co-infection with either adenovirus or herpes virus for its growth and multiplication, able of providing helper functions. A variety of serotypes are known, and contemplated herein. Such serotypes include, but are not limited to:

AAV-1 (Genbank accession no. NC002077; accession no. VR-645); AAV-2 (Genbank accession no. NC001401; accession no. VR-680); AAV-3 (Genbank accession no. NC001729; accession no. VR-681); AAV-3b (Genbank accession no. NC001863); AAV-4 (Genbank accession no. NC001829; ATCC accession no. VR-646); AAV-6 (Genbank accession no.NC001862); and avian associated adeno-virus (ATCC accession no. VR-1449). The preparation and use of AAVs as vectors for gene expression *in vitro* and for *in vivo* use for gene therapy are well known (see, *e.g.*, U.S. Patent Nos. 4,797,368, 5,139,941, 5,798,390 and 6,127,175; Tessier *et al.* (2001) *J. Virol. 75:*375-383; Salvetti *et al.* (1998) *Hum Gene Ther 20*:695-706; Chadeuf *et al.* (2000) *J Gene Med 2*:260-268).

Please replace the paragraph on page 10, lines 27-31, with the following:

As used herein, ϵ(efficiency), is the slope at the inflection point of the Hill curve (or, in general, of any other sigmoidal or linear approximation), to assess the efficiency of the global reaction (the biological agent and the assay system taken together) to elicit the biological or pharmacological response.

Please replace the paragraph on page 11, lines 9-13, with the following:

As used herein, a library of mutants refers to a collection of plasmids or other vehicles that carry (encode) the gene variants, such that individual plasmids or other vehicles carry individual gene variants. When a library of proteins is contemplated, it will be so-stated.

Please replace the paragraph on page 11, lines 16-23,, with the following:

As used herein, "reporter" or "reporter moiety" refers to any moiety that allows for the detection of a molecule of interest, such as a protein expressed

by a cell. Reporter moieties include, but are not limited to, for example, fluorescent proteins, such as red, blue and green fluorescent proteins; lacZ and other detectable proteins and gene products. For expression in cells, nucleic acid encoding the reporter moiety can be expressed as a fusion protein with a protein of interest or under the control of a promoter of interest.

a8

Please replace the paragraph on page 12, lines 1-7, with the following:

As used herein, activity refers to the function or property to be evolved. An active site refers to a site(s) responsible or that participates in conferring the activity or function. The activity or active site evolved (the function or property and the site conferring or participating in conferring the activity) may have nothing to do with natural activities of a protein. For example, it could be an 'active site' for conferring immunogenicity (immunogenic sites or epitopes) on a protein.

a9

Please replace the paragraphs beginning on page 14, line 35, through page 16, line 2, with the following:

As used herein, nucleic acids include DNA, RNA and analogs thereof, including protein nucleic acids (PNA) and mixture thereof. Nucleic acids can be single or double stranded. When referring to probes or primers, optionally labeled, with a detectable label, such as a fluorescent or radiolabel, single-stranded molecules are contemplated. Such molecules are typically of a length such that they are statistically unique of low copy number (typically less than 5, preferably less than 3) for probing or priming a library. Generally a probe or primer contains at least 14, 16 or 30 contiguous of sequence complementary to or identical to a gene of interest. Probes and primers can be 10, 14, 16, 20, 30, 50, 100 or more nucleic acid bases long.

a10

As used herein, homologous means about greater than 25% nucleic acid sequence identity, preferably 25% 40%, 60%, 80%, 90% or 95%. The intended percentage will be specified. The terms "homology" and "identity" are often used interchangeably. In general, sequences are aligned so that the

highest order match is obtained (see, *e.g.*: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; Carillo *et al.* (1988) *SIAM J Applied Math 48*:1073). By sequence identity, the number of conserved amino acids are determined by standard alignment algorithms programs, and are used with default gap penalties established by each supplier. Substantially homologous nucleic acid molecules would hybridize typically at moderate stringency or at high stringency all along the length of the nucleic acid of interest. Also contemplated are nucleic acid molecules that contain degenerate codons in place of codons in the hybridizing nucleic acid molecule.

Q10

As used herein, a nucleic acid homolog refers to a nucleic acid that includes a preselected conserved nucleotide sequence, such as a sequence encoding a therapeutic polypeptide. By the term "substantially homologous" it is meant having at least 80%, preferably at least 90%, most preferably at least 95% homology therewith or a less percentage of homology or identity and conserved biological activity or function.

Please replace the paragraph beginning on page 18, line 26, through page 19, line 17, with the following:

As used herein, genetic therapy involves the transfer of heterologous nucleic acids to the certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The nucleic acid, such as DNA, is introduced into the selected target cells in a manner such that the heterologous nucleic acid, such as DNA, is expressed and a therapeutic product encoded thereby is produced. Alternatively, the

Q19

heterologous nucleic acid, such as DNA, may in some manner mediate expression of DNA that encodes the therapeutic product, or it may encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy may also be used to deliver nucleic acid encoding a gene product that replaces a defective gene or supplements a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid may encode a therapeutic compound, such as a growth factor or inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous nucleic acid, such as DNA, encoding the therapeutic product may be modified prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof. Genetic therapy may also involve delivery of an inhibitor or repressor or other modulator of gene expression.

Q19

Please replace the paragraph on page 20, lines 16-18, with the following:

Q20

As used herein, a therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of disease.

Please replace the paragraph on page 21, lines 4-11, with the following:

Thus, by "isolated" it is meant that the nucleic is free of the coding sequences of those genes that, in the naturally-occurring genome of the organism (if any), immediately flank the gene encoding the nucleic acid of interest. Isolated DNA may be single-stranded or double-stranded, and may be genomic DNA, cDNA, recombinant hybrid DNA, or synthetic DNA. It may be identical to a native DNA sequence, or may differ from such sequence by the deletion, addition, or substitution of one or more nucleotides.

Q21

Please replace the paragraph on page 22, lines 19-27, with the following:

Q22

As used herein, the phrase "operatively linked" generally means the sequences or segments have been covalently joined into one piece of DNA,

whether in single or double stranded form, whereby control or regulatory sequences on one segment control or permit expression or replication or other such control of other segments. The two segments are not necessarily contiguous. For gene expression a DNA sequence and a regulatory sequence(s) are connected in such a way to control or permit gene expression when the appropriate molecules, e.g., transcriptional activator proteins, are bound to the regulatory sequence(s).

R22

Please replace the paragraph on page 23, lines 15-28, with the following:

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form, are not bound to the chromosome. "Plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. Other such forms of expression vectors that serve equivalent functions and that become known in the art can be used subsequently hereto.

R23

Please replace the paragraph beginning on page 25, line 25, through page 26, line 6, with the following:

As used herein, an array refers to a collection of elements, such as nucleic acid molecules, containing three or more members. An addressable array is one in which the members of the array are identifiable, typically by position on a solid phase support or by virtue of an identifiable or detectable label, such as by color, fluorescence, electronic signal (*i.e.* RF, microwave or other frequency that does not substantially alter the interaction of the molecules

R24

of interest), bar code or other symbology, chemical or other such label. Hence, in general the members of the array are immobilized to discrete identifiable loci on the surface of a solid phase or directly or indirectly linked to or otherwise associated with the identifiable label, such as affixed to a microsphere or other particulate support (herein referred to as beads) and suspended in solution or spread out on a surface.

R24

Please replace the paragraphs beginning on page 27, line 14, through page 29, line 9, with the following:

Recombinant viruses have been developed for use as gene therapy vectors. Gene therapy applications are hampered by the need for development of vectors with traits optimized for this application. The high throughput methods provided herein are ideally suited for development of such vectors. In addition to use for development of recombinant viral vectors for gene therapy, these methods can also be used to study and modify the viral vector backbone architecture, trans-complementing helper functions, where appropriate, regulatable and tissue specific promoters and transgene and genomic sequence analyses. Recombinant AAV (rAAV) is a gene therapy vector that can serve these and other purposes.

R25

The rep protein is an adeno-associated virus protein involved in a number of biological processes necessary to AAV replication. The production of the rRep proteins enables viral DNA to replicate, encapsulate and integrate (McCarty *et al.* (1992) *J. Virol 66*:4050-4057; Horer *et al.* (1995) *J. Virol 69*:5485-5496, Berns *et al.* (1996) Biology of Adeno-associated virus, in Adeno-associated virus (AAV) Vectors in Gene Therapy, K.I. Berns and C. Giraud, Springer (1996); and Chiorini *et al.* (1996) The Roles of AAV Rep Proteins in gene Expression and Targeted Integration, *from* Adeno-associated virus (AAV) Vectors in Gene Therapy, K.I. Berns and C. Giraud, Springer (1996)). A rep protein with improved activity could lead to increased amounts of virus progeny thus allowing higher productivity of rAAV vectors.

Since the Rep protein is involved in replication it can serve as a target for increasing viral production. Since it has a variety of functions and its role in replication is complex, it has heretofore been difficult to identify mutations that result in increase viral production. The methods herein, which rely on *in vivo* screening methods, permit optimization of its activities as assessed by increases in viral production. Provided herein are Rep proteins and viruses and viral vectors containing the mutated Rep proteins that provide such increase. The amino acid positions on the rep proteins that are relevant for rep proteins activities in terms of AAV or rAAV virus production are provided. Those amino acid positions are such that a change in the amino acid leads to a change in protein activity either to lower activity or increase activity. As shown herein, the alanine or amino acid scan revealed the amino acid positions important for such activity (i.e. hits). Subsequent mutations produced by systematically replacing the amino acids at the hit positions with the remaining 18 amino acids produced so-called "leads" that have amino acid changes and result in higher virus production. In this particular example, the method used included the following specific steps.

Amino acid scan

In order to first identify those amino acid (aa) positions on the rep protein that are involved in rep protein activity, an Ala-scan was performed on the rep sequence. For this, each aa in the rep protein sequence was individually changed to Alanine. Any other amino acid, particularly another amino acid such as Gly or Ser that has a neutral effect on structure, could have been used. Each resulting mutant rep protein was then expressed and the amount of virus it produced was measured. The relative activity of each individual mutant compared to the native protein is indicated in FIG 2A. HITS are those mutants that produce a decrease in the activity of the protein (in the example: all the mutants with activities below about 20 % of the native activity).

A25

R25

In a second experimental round, which included a new set of mutations and phenotypic analysis, each amino acid position hit by the Ala-scan step, was mutated by amino acid replacement of the native amino acid by the remaining 18 amino acids, using site-directed mutagenesis.

Please replace the paragraph on page ~~19~~ 29, lines 14-22, with the following:

A plasmid library was thus generated in which each plasmid contained a different mutant bearing a different amino acid at a different hit position. Again, each resulting mutant rep protein was then expressed and the amount of virus it could produce measured as indicated below. The relative activity of each individual mutant compared to the native protein is indicated in FIGURE 2B. Leads are those mutants that lead to an increase in the activity of the protein (in the example: the ten mutants with activities higher, typically between 2 to 10 times or more, generally 6-10 time, than the native activity).

R26

Please replace the paragraph beginning on page 30, line 21, through page 31, line 8, with the following:

The number of infectious particles produced was determined in a cell-based assay in which the activity of a reporter gene, in the exemplified embodiment, the bacterial lacZ gene, or virus replication (Real time PCR) was performed to quantitatively assess the number of viruses. The limiting dilution (titer) for each virus preparation (each coming from a different rep variant) was determined by serial dilution of the viruses produced, followed by infection of appropriate cells (293 HEK or HeLa rep/cap 32 cells) with each dilution for each virus and then by measurement of the activity of the reporter gene for each dilution of each virus. Hill plots (NAUTSCAN™) (published as International PCT application No. WO 01/44809 based on PCT n° PCT/FR00/03503, Dec, 2000; see EXAMPLES) or a second order polynomial function (Drittanti *et al.* (2000) *Gene Ther. 7*: 924-929; see co-pending U.S. Provisional Application 60/315,382) was used to analyze the readout data and to calculate the virus titers. Briefly, the titer was calculated from the second order polynomial

R27

function by non-linear regression fitting of the experimental data. The point where the polynomial curve reaches its minimum is considered to be the titer of the rAAV preparation. Results are shown in the EXAMPLE below.

Please replace the paragraph on page 32, lines 3-17, with the following:

In particular, as described in the Example, mutations in the Rep-encoding region of AAV, including serotypes AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 and AAV-6 are provided (see Example below). The mutant proteins and mutant adeno-associate virus (AAV) Rep proteins are provided. Exemplary proteins with mutations at one or more of residues 4, 20, 22, 29, 32, 38, 39, 54, 59, 124, 125, 127, 132, 140, 161, 163, 193, 196, 197, 221, 228, 231, 234, 258, 260, 263, 264, 334, 335, 337, 342, 347, 350, 354, 363, 364, 367, 370, 376, 381, 389, 407, 411, 414, 420, 421, 422, 424, 428, 438, 440, 451, 460, 462, 484, 488, 495, 497, 498, 499, 503, 511, 512, 516, 517, 518, 542, 548, 598, 600 and 601 of AAV-2 or the corresponding residues in other serotypes are provided. Residue 1 corresponds to residue 1 of the Rep78 protein encoded by nucleotides 321-323 of the AAV-2 genome (see Figure 3 and the Table below for an alignment of the mutations from various serotypes).

Please replace the paragraphs beginning on page 33, line 21, through page 34, line 8, with the following:

The rAAV provided herein are intended for use as vectors for gene therapy. The rAAV provided herein are intended for use in any gene therapy protocol the uses AAV as a vector. The mutant Rep proteins and nucleic acid molecules can be used to replace the corresponding gene in other AAV vectors. Of interest are the mutations provided herein that increase rAAV production. In particular, the mutant Rep proteins are used to increase production of rAAV derived from any of the AAV serotypes, including AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 and AAV-6 serotypes.

Toxicity and therapeutic efficacy of the rAAV can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LDS_{50} (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Doses that exhibit large therapeutic indices are preferred. Doses that exhibit toxic side effects may be used; care should be taken to design a delivery system that targets rAAV to the site of treatment in order to minimize damage to untreated cells and reduce side effects.

Please replace the paragraph beginning on page 34, line 24, through page 35, line 4, with the following:

AAV, which is a helper-dependent parvovirus requires co-infection with an adenovirus, herpes virus or papilloma virus (PV) for replication and particle formation. AAV inhibits PV-induced oncogenic tansformation, and this inhibition has been mapped to the Rep78 protein. The Rep78 protein inhibits expression of the PV promoter just upstream of the E6 gene (p89 of bovine PV-1 (BPV-1)) p97 of human PV-16 (HPV-16), and p105 of human PV-18 (HPV-18)). DNA binding is required for this inhibition. Rep78 also binds to the TAR sequences (nt +23 to +42) and to a region just upstream of the TATA box (nt. -54 to -34) in the HIV LTR region. AAV Rep78 also regulates a variety of other cancer associated genes, including, but are not limited to, C-H-*ras* (Khleif *et al.* (1991) *Virology 181*:738-741), c-*fos* and c-*myc* (Hermonat (1994) *Cancer Lttrs 81*:129-136).

Please replace the paragraphs beginning on page 35, line 12, through page 36, line 13, with the following:

Contemplated herein are AAV rep mutants that bind with greater affinity than wild-type AAV Rep78 to nucleic acid from PV, AAV, oncogenes or HIV, particularly HIV-1, and particularly promoter and other transcriptional/translational regulatory sequences from these sources. The mutant Rep protein when administered to a subject can inhibit PV and PV-associated diseases, HIV and HIV-associated diseases. Hence methods for treatment of PV and HIV-mediated disorders by administration of rAAV encoding mutant the Rep78 genes are provided. The particular mutants for use in these methods can be identified by testing each mutant for inhibitory activity, for example, in cell-based assays. For example, the Rep mutant protein can be tested by contacting it with nucleic acid from a PV, AAV or HIV or oncogene for a time sufficient to permit binding thereto, and comparing such binding to the binding of a wild-type Rep protein under the same conditions. Alternatively competitive binding assays may be performed. Mutant proteins having higher binding affinities are identified.

Fusion proteins containing a *tat* protein of HIV or other targeting agent and mutant Rep protein are also provided. Pharmaceutical compositions containing such fusion proteins are provided. The fusion proteins can contain additional components, such as *E. coli* maltose binding protein (MBP) that aid in uptake of the protein by cells (see, International PCT application No. WO 01/32711). Nucleic acid molecules encoding the mutant Rep protein or fusion protein operably linked to a promoter, such as an inducible promoter for expression in mammalian cells are also provided. Such promoters include, but are not limited to, CMV and SV40 promoters; adenovirus promoters, such as the E2 gene promoter, which is responsive to the HPV E7 oncoprotein; a PV promoter, such as the PBV p89 promoter that is responsive to the PV E2 protein; and other promoters that are activated by the HIV or PV or oncogenes.

A31

The mutant rep proteins are also delivered to the cells in rAAV or a portion thereof that can additionally encode therapeutic agents for treatment of the cancer or HIV infection or other disorders.

Please replace the paragraphs beginning on page 36, line 21, through page 37, line 27, with the following:

Other systems

It has been shown that the Rep protein is involved in the regulation of gene expression, including viral replication as described above, cellular pathways and protein phosphorylation (see, *e.g.*, Chiorini *et al.* (1998) *Mol. Cell Biol. 18*:5921-5929). Hence the mutant Rep proteins provided herein can be used to block, stimulate, inhibit, regulate or otherwise modulate metabolic or cellular signaling pathways. Rep proteins provided herein can be used to block, stimulate, inhibit, regulate or otherwise modulate cyclic AMP response pathways, and also to regulate or modulate cellular promoters as a means of modulating gene expression. Methods using these proteins for such purposes are provided herein.

A32

Formulation of rAAV

Pharmaceutical compositions containing the rAAV, fusion proteins or encoding nucleic acid molecules can be formulated in any conventional manner by mixing a selected amount of rAAV with one or more physiologically acceptable carriers or excipients. For example, the rAAV may be suspended in a carrier such as PBS (phosphate buffered saline). The active compounds can be administered by any appropriate route, for example, orally, parenterally, intravenously, intradermally, subcutaneously, or topically, in liquid, semi-liquid or solid form and are formulated in a manner suitable for each route of administration. Preferred modes of administration include oral and parenteral modes of administration.

The rAAV and physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or for oral, buccal, parenteral or rectal administration. For administration by inhalation, the rAAV can be delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g. dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges, e.g. of gelatin, for use in an inhaler or insufflator may be formulated containing a powder mix of a therapeutic compound and a suitable powder base such as lactose or starch.

A32

Please replace the paragraph on page 40, lines 1-4, with the following:

A33

The following example is included for illustrative purposes only and is not intended to limit the scope of the invention. The specific methods exemplified can be practiced with other species. The examples are intended to exemplify generic processes.

Please replace the paragraphs beginning on page 40, lines 13, through page 41, line 5, with the following:

Plasmids:

pNB-Adeno, which encodes the entire E2A and E4 regions and VA RNA I and II genes of Adenovirus type 5, was constructed by ligating into the polylinker of multiple cloning site of pBSII KS (+/-) (Stratagene, San Diego, USA) the SalI-HindIII fragment (9842-11555 nt) of Adenovirus type 5 and the BamHI-ClaI fragment (21563- 35950) of pBR325. All fragments of adenovirus gene were obtained from the plasmid pBHG-10 (Microbix, Ontario, Canada). pNB-AAV encodes the genes rep and cap of AAV-2 and was constructed by ligation of XbaI-XbaI PCR fragment containing the genome of AAV-2 from

A34

nucleotide 200 to 4480 into XbaI site of polylinker MCS of pBSIIKS(+/-). The PCR fragment was obtained from pAV1 (ATCC, USA). Plasmid pNB-AAV was derived from plasmid pVA1I, which contains the AAV genomic region, rep and cap. pNB-AAV does not contain the AAV ITR's present in pAV1. pAAV-CMV(nls)LacZ was provided by Dr Anna Salvetti (CHU, Nantes).

Plasmid pCMV(nls)LacZ (rAAV vector plasmid) and pNB-Adeno were prepared in DH5a E.coli and purified by Nucleobond AX PC500 Kit (Macherey-Nagel), according to standard procedures. Plasmid pAAV-CMV(nls)LacZ is derived from plasmid psub201 by deleting the rep-cap region with SnaB I and replacing it with an expression cassette harboring the cytomegalovirus (CMV) immediate early promoter (407 bp), the nuclear localized β-galactosidase gene and the bovine growth hormone polyA signal (324 bp) (see, Chadeuf *et al.* (2000) *J. Gene Med. 2*:260-268. pAAV-CMV(nls)LacZ was provided by Dr Anna Salvetti.

A34

Please replace the paragraph on page 43, lines 9-22, with the following:

rAAV from each of the above wells, were produced by triple transfection on 293 HEK cells. 3 x 10^4 cells were seeded into each well of 96 micro-well plate and cultured for 24 hours before transfection. Transfection was made on cells at about 70% confluency. 25 kDa PEI (poly-ethylene-imine, Sigma-Aldrich) was used for the triple transfection step. Equimolar amounts of the three plasmids AV helper plasmid (pNB-Adeno), AAV helper plasmid (pNB-AAV or a mutant clone rep plasmid) and vector plasmid (pAAV-CMV(nls)LacZ) were mixed with 10 mM PEI by gently shaking. The mixture was the added to the medium culture on the cells. 60 hours after transfection, the culture medium was replaced with 100 μl of lysis buffer (50mM Hepes, pH 7.4; 150 mM NaCl; 1mM $MgCl_2$; 1 mM $CaCl_2$; 0.01% CHAPS). After one cycle of freeze-thawing the

A35

A35

cellular lysate was filtered through a millipore filter 96 well plate and stored at −80°C.

Please replace the paragraph on page 44, lines 17-22, with the following:

A36

After 48 hours of infection, cells were treated with trypsin, and 100 μl of reaction solution (GalScreen Kit, Tropix) was added and incubated for one hour at 26 °C. Luminescence was measured in NorthStar (Tropix) HTS station. dRA plots were obtained plotting the intensity of β-Galatosidase activity vs. the dilution of the rAAV preparation.

Please replace the paragraph on page 45, lines 3-20, with the following:

A37

To identify candidate amino acid (aa) positions on the rep protein involved in rep protein activity an Ala-scan was performed on the rep sequence. For this, each amino acid in the rep protein sequence was replaced with Alanine. To do this sets of rAAV that encode mutant rep proteins in which each differs from wild type by replacement of one amino acid with Ala, were generated. Each set of rAAV was individually introduced into cells in a well of a microtiter plate, under conditions for expression of the rep protein. The amount of virus that could be produced from each variant was measured as described below. Briefly, activity of Rep was assessed by determining the amount of AAV or rAAV produced using infection assays on HeLa Rep-cap 32 cells and by measurement of AAV DNA replication using Real Time PCR, or by assessing transgene (β-galactosidase) expression. The relative activity of each individual mutant compared to the native protein was assessed and "hits" identified. Hit positions are the positions in the mutant proteins that resulted in an alteration (selected to be at least about 20%), in this instance all resulted in a decrease, in the amount of virus produced compared to the activity of the native (wildtype) gene (see Fig. 2A).

Please replace the paragraph on page 52, lines 4-11, with the following:

Sets of nucleic acids encoding the rep protein were generated. The rep proteins encoded by these sets of nucleic acid molecules were those in which each amino acid position identified as a "hit" in the ala-scan step, were each sequentially replaced by all remaining 18 amino acids using site directed mutagenesis. Each mutant was designed, generated, processed and analyzed physically separated from the others in addressable arrays. No mixtures, pools, nor combinatorial processing were used.

Please replace the paragraphs beginning on page 52, line 24, through page 53, line 12, with the following:

Based on the results obtained from the assays described above (i.e. titer of virus produced by each rep variant), each individual rep variant was assigned a specific activity. Those variant proteins displaying the highest titers were selected as leads (see Table above). Leads include: amino acid replacement of T by N at Hit position 350; T by I at Hit position 462; P by R at Hit position 497; P by L at Hit position 497; P by Y at Hit position 497; T by N at Hit position 517; L by S at Hit position 542; R by S at Hit position 548, G by S at Hit position 598; G by D at Hit position 598; V by P at Hit position 600.

Also provided are combinations of the above mutant Rep 78, 68, 52, 40 proteins, nucleic acids encoding the proteins, and recombinant AAV (any serotype) containing the mutation at the indicated position or corresponding position for serotypes other than AAV-2, including any set forth in the following table and corresponding SEQ ID Nos. Each amino acid sequence is set forth in a separate sequence ID listing; for each mutation or combination thereof there is a single SEQ ID setting forth the unspliced nucleic acid sequence for Rep78/68, which for all mutations from amino acid 228 on, includes the corresponding Rep 52 and Rep 40 encoding sequence as well.

Please replace the paragraph on page 72, lines 17-25, with the following:

Mutant adeno-associated virus (AAV) Rep proteins and viruses encoding such proteins that include mutations at one or more of residues 64, 74, 88, 175, 237, 250 and 429, where residue 1 corresponds to residue 1 of the Rep78 protein encoded by nucleotides 321-323 of the AAV-2 genome, and where the amino acids are replaced as follows: L by A at position 64; P by A at position 74; Y by A at position 88; Y by A at position 175; T by A at position 237; T by A at position 250; D by A at position 429 are provided. Nucleic acid molecules encoding these viruses and the mutant proteins are also provided.

Please replace the paragraph on page 73, lines 14-21, with the following:

For all of the mutant proteins provided herein those with increased activity, such as an increase in titer of rAAV when virus containing such mutations and/or expressing such mutant proteins are replicated, are of particular interest. Such mutations and proteins are provided herein and may be made by the methods herein, including by combining any of the mutations provided herein to produce additional mutant proteins that have altered biological activity, particularly increased activity, compared to the wild-type.

Please replace thé paragraph on page 74, lines 3-17, with the following:

Also provided are nucleic acid molecules and rAAV (any serotype) in which position 630 (or the corresponding position in another serotype; see Figs. 3 and the table above) has been changed. Changes at this position and the region around it lead to changes in the activity or in the quantities of the Rep or Cap proteins and/or the amount of AAV or rAAV produced in cells transduced with AAV encoding such mutants. Such mutations include tgc to gcg change (SEQ ID No. 721). Mutations at any position surrounding the codon position 630 that increase or decrease the Rep or Cap proteins quantities or activities are also provided. Methods using the rAAV (any serotype) that contain nucleic